DOCUMENT RESUME

ED 213 026 CS 206 742

AUTHOR Felker, Daniel B.; Rose, Andrew M.

TITLE The Evaluation of a Public Document: The Case of FCC's Marine Radio Rules for Recreational Boaters. Document Design Project, Technical Report No. 11.

INSTITUTION American Institutes for Research, Washington, D.C.; Carnegie-Mellon Univ., Pittsburgh, Pa.; Siegel & Gale, Inc., New York, N.Y.

SPONS AGENCY National Inst. of Education (ED), Washington, DC. Teaching and Learning Program.
REPORT NO AIR-75003-1/81-TR
PUB DATE Feb 81
CONTRACT 400-78-0043
NOTE 42p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adults; Comparative Analysis; Content Analysis; *Government Publications; *Language Research; *Language Usage; *Layout (Publications); *Readability
IDENTIFIERS Document Design Project; Federal Communications Commission

ABSTRACT

In a collaborative effort, the Federal Communications Commission (FCC) and the Document Design Project conducted an evaluation of marine radio rules for recreational boaters that had been rewritten in plain English by FCC personnel. The revised rules were evaluated by 53 experienced boaters and 52 inexperienced boaters, who were given either the old rules or the new ones and then tested for their comprehension of the rules. Each question on the test required the subject to find the appropriate rule and apply its provisions. The subjects also noted the time it took them to complete each question and rated the overall difficulty they had in using the rules. Results showed that subjects using the revised rules were significantly better at identifying the proper rules, took less time to answer questions about the rules, and rated them as easier to use than did subjects using the old rules. (A copy of the comprehension test is appended.) (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED213026

AIR-75003-1/81-TR

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

DOCUMENT DESIGN PROJECT

Technical Report No. 11

The Evaluation of a Public Document:

The Case of FCC's Marine Radio Rules for Recreational Boaters

Daniel B. Felker

Andrew M. Rose

A Product of the Document Design Project funded by the National Institute of Education, Washington, D.C.
Contract No. NIE-400-78-0043

February 1981

CS206742

AMERICAN INSTITUTES FOR RESEARCH/1055 Thomas Jefferson Street, NW, Washington, DC 20007 • 202/342-5000
with Carnegie Mellon University and Siegel & Gale, Inc.

This report is part of the work being done under contract #400-78-0043 of the National Institute of Education. The project officer is Dr. Candace Miyamura of the Teaching and Learning/ Reading and Language Group. This report does not necessarily reflect the views of the funding agency.

THE EVALUATION OF A PUBLIC DOCUMENT:

The Case of FCC's Marine Radio Rules for Recreational Boaters

Abstract

The Federal Communications Commission revised its marine radio rules for recreational boaters by writing it in plain English and by incorporating various document design principles. The revised rules were evaluated by a 2 X 2 factorial ANOVA. One hundred five subjects (formed into experienced and non-experienced boater groups) were compared on how well they used and understood the original and revised rules. Groups using the revised rules were significantly better in identifying the proper rules. In answering questions about the rules, they took less time to answer questions. They also rated the revised rules as easier to use. These results and the implications of conducting empirical evaluations of public documents are discussed.

Introduction

Background

The Federal Communications Commission (FCC) is the agency concerned with regulating all non-government wire and radio communications. The FCC's sphere of influence and responsibility is immense: it oversees all public and commercial TV and radio broadcasts; it regulates interstate telephone rates; and it supervises all two-way radio services used by police and fire departments and used in aviation and marine operations. In carrying out its mandate, the FCC publishes, monitors, and enforces

many thousands of rules and regulations aimed at audiences ranging from lawyers and engineers to the truck driver buying a CB radio. And, in common with other government agencies, most of the rules and regulations issued by the FCC are hard to read and understand.

The FCC has been in the forefront of Federal agencies to take serious steps to simplify some of its rules and regulations. The rules governing citizens band radio service (CB radios) were rewritten into plain English in 1978. The lawyers who rewrote these rules completely reorganized them. They replaced technical words and legal phrases with common terms. They used informative headings and tables instead of dense prose. White space was liberally used between paragraphs and in the margins. The CB rules were widely recognized throughout the Federal establishment as a model for others in the government to follow in order to make public rules and regulations easier to read and understand. However, no empirical evidence was ever offered to support the claim that the revised rules are comprehensible or that they are an improvement over the original version.

In a second major effort, the FCC revised Part 83 of Chapter I of Title 47 of the Code of Federal Regulations, which deals with FCC's marine radio rules for recreational boaters. The FCC's goal was to produce a low cost booklet of rules that the 300,000 recreational boaters who have two-way marine radios on their boats could use and understand. The writers extracted information that specifically relates to recreational boating from a much larger set of rules which also covered radio communication among large

commercial ships and vessels. In revising the marine radio rules into an 11-page booklet, they followed document design principles that are similar to the ones that the authors of CB radio rules had used. This time around, the FCC wanted to know whether the revised rules, in fact, are usable, readable, and easy to understand, and whether they are any better than the original rules.

FCC staff directly involved in revising the rules contacted the Document Design Project (DDP) of the American Institutes for Research (AIR) for assistance in developing and conducting an empirical evaluation of the revised rules. The DDP agreed to provide technical assistance to FCC at no cost under the provisions of AIR's contract with the National Institute of Education. This contract permits AIR to support selected document design research and development efforts in Federal agencies as cases of putting research into practice.

Purpose

The evaluation was a collaborative effort between the FCC and the Document Design Project. Each group had different ends that they wanted the evaluation to achieve, although these ends were complementary.

One major and obvious goal of the FCC was to determine whether all the effort that went into rewriting the marine radio rules, while continuing to fulfill other day-to-day responsibilities, was worth it. Did revising the rules into plain English and did using

recommended document design principles result in an understandable, usable document? Was it any improvement over the existing rules?

The second goal of the FCC's designers was more implicit and political. They wanted to demonstrate to skeptical superiors that the resources expended for improving documents are worthwhile investments. It takes considerable money, time, and staff for government agencies to formulate, develop, and approve rules and regulations. Yet, if existing staff can be trained to design better public documents, and if the same staff can evaluate these documents and demonstrate that the public understands and is able to use them, then the investment costs become defensible.

For its part, the Document Design Project of AIR had both scientific and practical purposes in supporting this evaluation. As document design scientists, we wanted empirical verification of the value of widely recommended document design principles that were used in the revised marine radio rules. For all of the recommendations that are offered for making documents clear and useful by countless writing and design "experts," the fact is that there is minimal support for their assertions in the research literature. This evaluation would help fill a gap in our knowledge about the effects of document design principles on document users.

At a more applied level, we wanted to test the ability of non-research trained government employees to design and execute an empirical evaluation with minimal assistance from experienced evaluators. One of our more sobering observations in the 2 1/2 years of providing technical assistance to government agencies is

that the vast majority of regulations, forms, and rules issued annually by the government, even those that directly affect the well being of millions of people, are never tested beforehand on the intended audience. Sometimes this is because the Federal agency misunderstands or is unaware of the concept of evaluation; in other cases, the Federal agency believes evalution is too complicated or too expensive to do. Because we firmly believe that critical public documents must be evaluated if they are to serve people, and because the resources available to support document evaluation are likely to remain limited, we were interested in the degree to which typical, non-research trained document designers can plan and execute useful, if not scientifically impeccable, empirical evaluations.

In the remainder of this report we will describe the evaluation along the lines of a conventional research report. However, unlike a typical research report, we will intersperse remarks that relate to practical issues of getting useful evaluations done in the real world of document design.

The Evaluation Plan

The evaluation paradigm we used was a straightforward comparative, experimental design. Basically, the evaluation experiment consisted of presenting the subjects, which in this case were experienced and inexperienced recreational boaters, with either the new or old marine radio rules. We then asked them 13 test questions that concerned information they had read in the

rules. We compared the performance of the subjects on this test in order to detect any differences between the new and old rules. With this overview in mind, we now describe the elements of the evaluation plan in more detail.

The Sample of Users

The proper subjects for this evaluation are the kinds of people who would be expected to use the marine radio rules while boating. This breaks down into two distinct audiences: experienced recreational boaters and inexperienced boaters. Moreover, given the constraints in resources and time, the subjects had to be drawn from populations convenient to FCC's document designers.

Experienced boaters were needed because the revised marine radio rules will replace the existing rules and we wanted to test whether familiarity with the old rules helps or hinders understanding the new rules. Inexperienced boaters were needed because new owners of recreational boats also will be expected to use the new rules. We could not presume that all boaters have prior experience with either marine radios or the existing rules so we wanted to test how well inexperienced boaters understand the new FCC rules.

The experienced subjects were 53 male members of a Power Squadron in the Washington, D.C. metropolitan area who volunteered to participate in the evaluation. The Power Squadron is an organization of boating enthusiasts who own their own recreational

boats and have extensive boating and marine radio experience. We refer to these subjects as the "Power" group.

The inexperienced subjects were 52 people recently employed by the FCC Licensing Division at Gettysburg, Pennyslvania. This group was predominantly female (48 out of 52). None had any experience with marine radios or with recreational boating. For convenience, we refer to these subjects as the "Gettysburg" group.

The sample thus consisted of 105 intended and potential users of the revised marine radio rules. The sample admittedly failed to meet the established norms of scientific research--e.g., the subjects were not randomly drawn from their respective populations, there was an uneven distribution of males and females. Nevertheless, we believe that the makeup of these groups was sufficiently representative of the potential audiences expected to use the marine radio rules. The sample was suitable for meeting the essential requirements of the evaluation.

The Materials

Old Rules. The "old" rules consisted of 49 pages that contained information related to the use of marine radios in recreational boats. These pages were taken from a larger volume of Part 83 of the original regulations. An index was specifically created to help the subjects locate particular rules in the 49 page document. A sample page of the old rules, written in typical "bureaucratese," is shown on the next page.

83.115 Retention of radio station logs.

(a) All station logs which are required under those provisions of this part pertaining to the particular classes of stations subject to this part shall be retained by the licensee for a period of one year from date of entry and for such additional periods as required by the following subparagraphs:

(1) Station logs involving communications incident to a distress or disaster shall be retained by the station licensee for a period of 3 years from date of entry;

(2) Station logs which include entries of communications incident to or involved in an investigation by the Commission and concerning which the station licensee has been notified shall be retained by the station licensee until such licensee is specifically authorized in writing by the Commission to destroy them;

(3) Station logs incident to or involved in any claim or complaint of which the station licensee has notice shall be retained by such licensee until such claim or complaint has been fully satisfied or until the same has been barred by statute limiting the time for the filing of suits upon such claims.

Note: See Part 42 of this chapter concerning preservation of records of common carriers.

(b) Station logs shall be made available to an authorized representative of the Commission upon request.

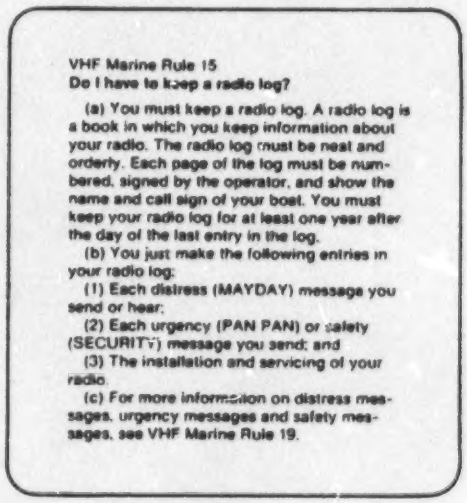
(c) Ship station logs shall be fully completed at the end of each voyage and before the operator(s) (or other person(s) responsible under the applicable provisions of this part) leave the ship. Unless otherwise authorized by the applicable provisions of this part, the radio log currently in use shall be kept by the licensed operator(s) of the station and during use shall be located at the principal radio operating room of the vessel. At the conclusion of each ocean voyage terminating at a port of the United States (includes Puerto Rico, and Virgin Islands), the original radio log (or a duplicate thereof) dating from the last departure of the vessel from a U.S. port shall be retained under proper custody on board the vessel for a suffcient period of time (not more than 24 hours) to be available for inspection by duly authorized representatives of the Commission. After retention on board the vessel as herein stipulated, the original log (and the duplicate log if provided) may be filed at an established shore office of the station licensee, and shall be retained as stipulated by paragraph (a) of this section.

Note: Duplicate logs are not required by the provisions of this paragraph, unless the original log is removed prior to opportunity for official inspection.

(d) Logs of ships of the United States containing entries required to be made by reason of the Great Lakes Agreement or 83.368(c) of this part shall be kept at the principal radiotelephone operating location while the vessel is being navigated. All entries in their original form required by said agreement or 83.368(c) shall be retained on board the vessel for a period od not less than one month from the date of entry. After retention on board the vessel as herein stipulated, the entries shall be filed at a place where they will be commission upon request, and shall be retained as stipulated by paragraph (a) of this section.

Example of "Old" Marine Radio Rule

New Rules. The revised rules that were developed by the FCC consisted of an 11-page booklet organized into 22 rules. The new rules covered the same content areas as the old rules, but were completely reorganized, redesigned, and rewritten. A sample rule of the new version is shown below. Note that paragraph 83.115 from the old rules and rule 15 from the new rules cover the same issues (keeping radio logs).



VHF Marine Rule 15
Do I have to keep a radio log?

(a) You must keep a radio log. A radio log is a book in which you keep information about your radio. The radio log must be neat and orderly. Each page of the log must be numbered, signed by the operator, and show the name and call sign of your boat. You must keep your radio log for at least one year after the day of the last entry in the log.

(b) You just make the following entries in your radio log:

(1) Each distress (MAYDAY) message you send or hear;

(2) Each urgency (PAN PAN) or safety (SECURITY) message you send; and

(3) The installation and servicing of your radio.

(c) For more information on distress messages, urgency messages and safety messages, see VHF Marine Rule 19.

Sample Rule from FCC's Revised Marine Radio Rules for Recreational Boaters

Performance Test. All subjects were administered a test designed to measure their comprehension of the rules. The test consisted of 13 questions about various aspects of marine radio operations. The questions represented important and frequently occurring situations where the regulations ought to be consulted. Since subjects did not have to read all sections of the regulations to be able to answer the questions, the test was not exhaustive. However, we believe the test provided a wide range of typical situations. As such, the experimental results could be generalized

to any other sample of questions that might have been chosen. A sample test question is shown below.

> Thomas Miller is licensed to operate a marine radio aboard his recreational boat. He then sells his boat to a neighbor and tells him that the ship station license goes with the boat. Is Mr. Miller correct?
>
> Yes or No (circle one)
>
> Why or Why not?

Each test question required subjects to first find the appropriate rule pertaining to the question and then apply its provisions to the question. Several questions had more than one part. In all, there were 20 possible "points." Thus, a subject who answered all the questions completely and accurately scored 20 on the test; those who answered fewer questions or answered them less accurately received lower scores. The subjects also noted the time when they began each question and when they finished. The complete test is presented in the appendix.

Design and Analysis

The design of the experiment required all subjects to answer the same questions. The 2 X 2 factorial experimental design (2 levels of experience and 2 forms of rules) is depicted in Figure 1.

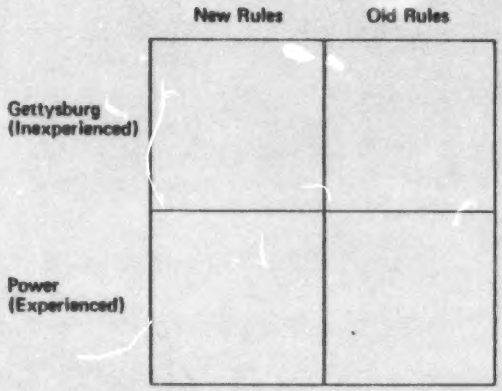


Figure 1. Experimental design used to evaluate FCC's revised rules for recreational boaters.

The Gettysburg and Power groups were tested separately, at different times and locations. Each subject was randomly assigned to either the new or old marine radio rules and given a copy of the 13-question test. The subjects answered each test question in sequence; they used the rules to note the particular rule that was appropriate for each question and also to answer the questions themselves. Each subject also recorded the time taken to complete each of the 13 questions. This was done by using wall-clock time, rounded off to the nearest minute. Finally, subjects rated the overall difficulty they had in using the marine radio rules.

In analyzing these data, we first described and summarized the results obtained for each of these measures. We then applied the analysis of variance (ANOVA) to determine the direction and degree of difference in the measures for each group. Finally, we conducted a "secondary" analysis that examined the patterns of errors made by each of the groups.

The rationale for these analyses rests on the relationships among the performance measures collected and the underlying concept that they are assumed to reflect. In this case, we assumed that the dependent measures--number of answers correct, number of rules correct, time per question, and ratings of difficulty--were valid reflections of the readability, understandability, and usefulness of the marine radio rules. The statistical tests we conducted were merely "checks" to determine the probability that any result could have occurred by chance.

Procedures

The subjects were run in large groups. After they were seated, each person was given a packet of material at random. The packet contained either the old or new versions of the rules and the test.

The FCC administrator explained the purpose of the study. The group was instructed to begin with question #1 and to note the particular rule that applied, answer the question by reading the appropriate rule, and enter the time taken to complete the question to the nearest minute. A large wall clock visible to the entire group was used for timing. Following question #13, the subjects rated the version of rules they used for difficulty. The subjects worked through the questions at their own pace. When the entire group had finished, the FCC administrator collected the rules and tests and answered any questions.

Results

As we noted earlier, the purpose of the evaluation was to ascertain whether the revised marine radio rules are comprehensible and usable, and to test whether they are an improvement over the original rules. To help us make these determinations, the evaluation generated four different scores for each subject:

- The number of answers correct (out of 20)--a measure of comprehensibility.
- The number of rules correctly identified (out of 13)--a measure of both usability and comprehensibility.
- The time needed to answer each question--a measure of usability.
- An overall rating of difficulty in using the rules--a measure of both usability and comprehensibility.

Table 1 on the next page presents the summary statistics of the results obtained from the experiment. We will explain Table 1, before proceeding to the specific analyses of each dependent measure.

Section "A" of Table 1 summarizes the number of test questions answered correctly. The first two columns illustrate the total number of possible correct answers and the actual number of correct answers for each group of subjects. For example, the first line of Section "A" shows that out of the 500 possible correct answers for the Gettysburg-Old group (i.e., 20 scorable test points X 25 subjects = 500), only 304 actually were answered correctly. The

TABLE 1
Basic Descriptive Information

A: Test Items Answered Correctly

Subject Group		Number of Possible Correct Answers	Number Answered Correctly	Mean Number Correct	Standard Deviation
Gettysburg - Old	(N=25)	500	304	8.54	3.21
Power - Old	(N=25)	500	422	12.78	3.78
Gettysburg - New	(N=27)	540	537	17.26	2.61
Power - New	(N=28)	560	549	16.45	3.53
Combined Old	(N=50)	1,000	726 (71%)	10.66	
Combined New	(N=55)	1,100	1,086 (99%)	16.85	

B: Rules Identified Correctly

Gettysburg - Old	(N=25)	325	203	6.60	2.36
Power - Old	(N=25)	325	238	6.96	3.23
Gettysburg - New	(N=27)	351	346	11.15	1.68
Power - New	(N=28)	364	336	9.80	3.40
Combined Old	(N=50)	650	441 (67%)	6.78	
Combined New	(N=55)	715	682 (95%)	10.47	

C: Time Per Question (Minutes)

		Number of Questions Answered	Mean Minutes to Answer Each Question	Standard Deviation
Gettysburg - Old	(N=25)	226	3.51	1.47
Power - Old	(N=25)	225	2.43	1.22
Gettysburg - New	(N=27)	322	1.73	0.63
Power - New	(N=28)	319	1.50	0.58
Combined Old	(N=50)	478	2.97	
Combined New	(N=55)	641	1.62	

D: Difficult Ratings (5=hard, 1=easy)

		Mean Difficulty Rating
Gettysburg - Old	(N=25)	4.57
Power - Old	(N=25)	4.22
Gettysburg - New	(N=27)	2.00
Power - New	(N=28)	1.77
Combined Old	(N=50)	4.39
Combined New	(N=55)	1.88

third column shows the mean (average) number of questions answered correctly per subject. The fourth column presents the standard deviation for each group. The major comparision of interest, the performance of subjects using the new rules vs. those using the old rules, is shown by the "Combined New and Old" headings.

Section "B" of Table 1 summarizes the number of rules correctly identified by the subjects and contains the same information as Section "A."

Section "C" of Table 1 summarizes the amount of time (in minutes) taken to answer the test items. The first column shows the total number of questions answered by each group. The second column presents the mean (average) number of minutes taken to answer each question for each group.

Section "D" presents the average rating of difficulty each group gave to the version of rules assigned to them (5=hard, 1=easy).

On the basis of Table 1 information alone, it is clear that there are performance differences between subjects using the old rules and those using the new rules. We discuss the statistical tests used to evaluate the significance of these differences next.

Questions Answered Correctly

Table 1 illustrates that subjects using the new rules answered more test questions correctly. Those using the new rules scored an average of 16.85 points (out of 20) compared to 10.66 points for

those using the old rules. Figure 2 graphically displays the differences in performance for each group. It is clear that both

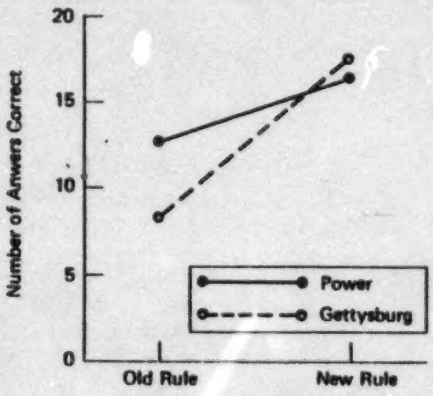


Figure 2. The average number of questions answered correctly (out of 20) by using old and new recreational boating rules.

experienced and inexperienced boaters answered more items corrrectly with the new rules. As might be expected, the experienced boaters performed better with the old rules (Power-Old) than inexperienced boaters (Gettysburg-Old), but this difference disappeared with the use of the new rules.

Results from the analysis of variance (ANOVA) indicated that:

- The difference between the old and new versions of the rules was statistically significant, $F(1,101) = 91.85$, $p<.01$.
- The difference between the two groups (Power and Gettysburg) was statistically significant, $F(1,101) = 7.03$, $p<.01$.
- The interaction of groups and rules was significant, $F(1,101) = 15.29$, $p<.01$.

Our interpretation of these results is straightforward. The

revised rules are significantly superior to the old rules in terms of answering test questions correctly. Moreover, inexperienced boaters are just as successful as experienced boaters (in fact, a little better) in their ability to use the new rules to answer the test questions. We conclude from this that the new rules are comprehensible.

Rules Identified Correctly

Figure 3 shows the average number of rules each group identified correctly when using the new or old rules. The subjects using the new rules correctly identified an average

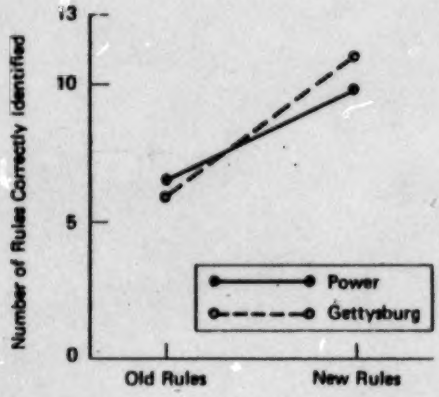


Figure 3. The average number of rules identified correctly (out of 13) by using old and new recreational boating rules.

of 10.47 rules (out of 13) and those using the old rules an average of 6.78. This difference is statistically significant, $F(1,101) = 46.85$, $p<.01$. Thus both experienced and inexperienced boaters were better able to locate and identify the correct rules needed for answering questions with the new rules. The groups did not differ in their ability to find the correct rule, and the interaction was

not significant. A reasonable interpretation of this finding is that the new rules are more comprehensible and easier to use.

Time to Answer

Figure 4 shows the average amount of time taken by the two groups to answer each test question using the old and new rules. Included in this is the time taken to read the questions, locate the proper rule, read the rule, and answer the questions.

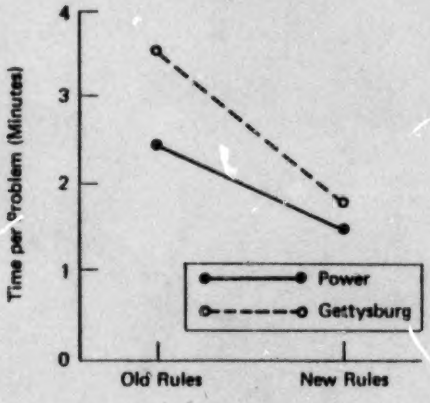


Figure 4. The average time taken to answer each test question using old and new rules.

Subjects using the new rules took significantly less time to answer test questions (1.62 minutes on the average vs. 2.97 minutes). In addition, The Power group was significantly faster overall and the interaction was significant (the Gettysburg group improved more than the Power group). These results were all predictable because the 49 pages of the old rules inherently are more cumbersome and time consuming to use than the 11 pages of the new rules. The Power groups could be expected to work with the old rules faster

due to prior experience. The results support the conclusion that the new rules are more usable for recreational boaters.

Difficulty Ratings

All subjects rated how difficult it was to use the particular set of rules that was assigned to them. The rating scale ranged from 5 (hard to use) to 1 (easy to use). Figure 5 shows the differences in the ratings given by the two groups.

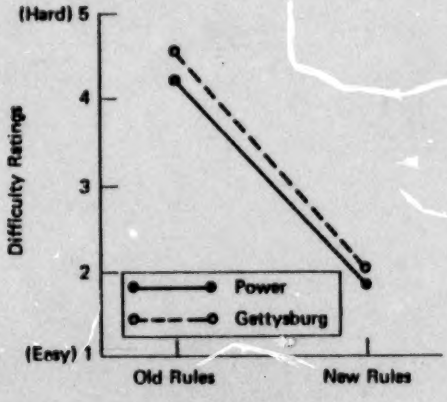


Figure 5. The average ratings of how difficult it was to use the new and old rules.

It is clear from the figure that both experienced and inexperienced boaters judged the new rules easier to use. This result was statistically significant and supports our interpretation that the new rules are more usable and more comprehensible.

Secondary Analysis: Errors

We conducted an additional analysis of the types of errors made by the different subject groups in an attempt to better

understand the differences in test performance. The subjects could make three types of errors on the test:

Type 1: The subject could answer a question correctly, yet give the wrong rule.

Type 2: The subject could answer a question incorrectly, but give the right rule.

Type 3: The subject could get both the question and the rule wrong.

Table 2 shows the proportion of errors of each type for the different subject groups. For example, of all errors made by the Gettysburg-New group, 30% were of the Type 1 (right answer, wrong rule), 38% were of the Type 2 (wrong answer, right rule), and 32% were of the Type 3 (wrong answer, wrong rule).

TABLE 2
Proportion and Type of Errors by Subject Groups

Subject Group	TYPE 1 ERROR (Right Answer, Wrong Rule)	TYPE 2 ERROR (Wrong Answer, Right Rule)	TYPE 3 ERROR (Wrong Answers, Wrong Rule)
Gettysburg - Old	.06 (11)*	.19 (38)	.75 (147)
Power - Old	.28 (53)	.12 (23)	.60 (114)
Gettysburg - New	.30 (24)	.38 (31)	.32 (26)
Power - New	.36 (43)	.24 (29)	.39 (47)

* Numbers in parentheses are the actual frequencies of each type of error.

Table 2 illustrates several interesting performance patterns. The first is the pattern of errors made by subjects who used the

old regulations. The Gettysburg-Old subjects made very few of the first type of error (right answer, wrong rule). This indicates that these inexperienced subjects really did not know much about marine radios. Furthermore, this finding adds support to the validity of the test in that subjects were not able to "guess" correct answers. In contrast, the Power-Old subjects, who were experienced boaters, were able to respond correctly, despite using the wrong rule. This, too, is an indicator that the test asked "good" questions because if experienced marine radio operators could not answer any questions, it would be possible to argue that the test was unrealistic.

Another interesting pattern shown in Table 2 is the large increase in proportion of Type 1 error (right answer, wrong rule) for the Gettysburg-New group as compared to the Gettysburg-Old group. A possible interpretation of this unexpected finding is that the Gettysburg-New subjects were able to read through the entire regulation and thereby increase their general knowledge of marine radios. Thus, they were able to answer some questions correctly despite going to the wrong rule.

A final pattern of interest in Table 2 is that Type 2 errors (wrong answer, right rule) still occurred with the new rules. This suggests that there are still some problems with the new rules, even with all the effort that went into writing them in plain English and in redesigning them.

Discussion

We pointed out earlier that both the FCC and the Document Design Project had two somewhat different goals for this evaluation. On the one hand, we both were interested in knowing whether government rules that were written in plain English and that incorporated other document design principles were understandable and usable, and whether the revised rules were any better than the original ones.

On the other hand, we both were concerned about the degree to which non-research trained document designers could carry out a useful evaluation of a public document. The FCC document designers were looking for evidence that demonstrated that they were capable of developing readable rules and regulations and that also could be used to convince their superiors of the value of document design activities. The DDP wanted to monitor the entire process of non-evaluators conducting an evaluation in order to find out whether they could do it and to learn more about how to help other non-research trained document designers conduct empirical evaluations.

Are the Revised Rules Clear and Usable?

The statistical evidence generated by the evaluation on all four dependent measures supports the conclusion that subjects using the new rules performed far better than subjects using the old rules. More questions were answered correctly, the correct rules

were identified more often, time per question was faster, and the new rules were judged easier to use. Moreover, these effects were found with both experienced and inexperienced recreational boaters. Even on measures where experienced boaters had an advantage by virtue of previous exposure to or knowledge of the old regulations, the advantage disappeared with the new rules.

The evaluation strongly suggests that the FCC group that revised the rules met their overall goal. The revised rules are indeed usable, readable, and easy to understand by the kinds of people expected to use them. In addition, they are superior to the existing rules. The use of simpler language, distinctive headings, logical presentation, and the elimination of technical jargon and superfluous information does make a difference. We do not know whether this increased clarity is cost-effective to the government in terms of the time and expense taken to revise the rule. Neither do we know if the new rules result in reduced errors and accidents by boaters, or if they resulted in increased good will of citizen boaters. This study did not examine these issues. Yet, it would be hard to argue against the value of a set of government rules that had the demonstrated effect on users that the FCC's marine radio rules had.

The implications are clear: government regulations can be improved and can be made more usable and comprehensible for their users. Similar results could occur by revising other cumbersome public documents, not only within the FCC but in other government agencies as well.

Is the Evaluation Useful?

The answer to this depends upon the answers to still further questions. Useful to whom? To the FCC? To DDP?

From the FCC's point of view the evaluation provides direct and consistent evidence that the revised marine radio rules are clear and usable to experienced and potential recreational boaters. This evidence is independent of any personal opinion about the quality of the revised rules or of any personal feelings about the value of the plain English movement. What the FCC does with the information generated by this evaluation is up to the revisors of the marine radio rules and the appropriate FCC decision makers. As document designers and evaluators, we would hope that the success of the effort both to rewrite and to evaluate rules will spur the FCC and other government agencies to conduct similar projects.

From the DDP's point of view the evaluation is a success. We believe that the experience has shown that non-research trained document designers can plan and conduct a useful empirical evaluation of a public document. We believe that typical document designers can learn how to carry out all essential evaluation tasks, with the possible exception of data analysis. They can learn to identify the purpose of an evaluation, to articulate the value of empirical evaluation, to identify the appropriate sample of document users and secure their cooperation, to design suitable performance measures, to implement standardized data collection activities, and to interpret the significance of any findings for future decisions.

APPENDIX

Performance Test

TIME BEGIN _____:_____

TIME FINISH _____:_____

1. How must you identify your station when your are communicating?

   (a) By the name of the vessel.

   (b) By the vessel's state registration or documentation number.

   (c) By FCC call sign.

   (d) By FCC call sign and boat name.

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH _____:_____

2. When must you identify your station?

   (a) At the beginning of a message.

   (b) At the end of a message.

   (c) At both the beginning and the end of a message.

   (d) Whenever you change channels and at the end of a message.

RULE SECTION: _____

TIME BEGIN _____:_____

TIME FINISH _____:_____

3. Do recreational boaters have to keep a radio log?

Yes or No (circle one)

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

4. What must you enter in a radio log?

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

5. Thomas Miller is licensed to operate a marine radio aboard his recreational boat. He then sells his boat to a neighbor and tells him that the ship station license goes with the boat. Is Mr. Miller correct?

   Yes or No (circle one)

   Why or Why Not?

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

6. If Thomas Miller's ship station license expired yesterday but he has applied for renewal, can he still operate his marine radio?
   Yes or No (circle one)

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

7. You are on a trip with several other boats from your yacht club and wish to remain in contact with one another. Since everyone's radio is equipped with channel 6, you agree to use this channel.

   Is this permitted?

   Yes or No (circle one)

   How about channel 68?

   Yes or No (circle one)

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

8. You are participating in an overnight cruise with several other recreational boats. Which of the following channels should you use to talk to one another:

   (a) Channel 12

   (b) Channel 68

   (c) Channel 22

   (d) Channel 6

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

9. Thomas Miller's marine radio has recently broken. Is it legal for Mr. Miller to repair his radio himself?

Yes or No (circle one)

If No, then who must repair it?

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH:_____:_____

10. Thomas Miller recently decided to put a marine radio on his boat. What FCC license or licenses does he need?

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

11. What form or forms should Mr. Miller use to apply for his license or licenses?

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

12. Thomas Miller is operating his recreational boat and sees a floating log which is a hazard to navigation. He wishes to contact the Coast Guard to report this hazard. What channel or channels should he use?

    (a) Channel 22

    (b) Channel 13

    (c) Channel 6

    (d) Channel 16

RULE SECTION: _____

TIME BEGIN: _____:_____

TIME FINISH: _____:_____

13. Which emergency signal would Thomas Miller use if he wished to warn another vessel of a navigational hazard?

(a) PAN PAN

(b) SECURITY

(c) MAYDAY

RULE SECTION: _____

14. Did you think these rules were-- (check one)

Very Difficult to Understand: _____

Difficult to Understand: _____

Of Average Difficulty: _____

Easy to Understand: _____

Very Easy to Understand: _____